



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

SARS-CoV-2 shedding sources in wastewater and implications for wastewater-based epidemiology

Xuan Li, Jagadeeshkumar Kulandaivelu, Ying Guo, Shuxin Zhang, Jiahua Shi, Jake O'Brien, Sudipti Arora, Manish Kumar, Samendra P. Sherchan, Ryo Honda, Greg Jackson, Stephen P. Luby, Guangming Jiang



PII: S0304-3894(22)00456-3

DOI: <https://doi.org/10.1016/j.jhazmat.2022.128667>

Reference: HAZMAT128667

To appear in: *Journal of Hazardous Materials*

Received date: 13 January 2022

Revised date: 5 March 2022

Accepted date: 8 March 2022

Please cite this article as: Xuan Li, Jagadeeshkumar Kulandaivelu, Ying Guo, Shuxin Zhang, Jiahua Shi, Jake O'Brien, Sudipti Arora, Manish Kumar, Samendra P. Sherchan, Ryo Honda, Greg Jackson, Stephen P. Luby and Guangming Jiang, SARS-CoV-2 shedding sources in wastewater and implications for wastewater-based epidemiology, *Journal of Hazardous Materials*, (2022) doi:<https://doi.org/10.1016/j.jhazmat.2022.128667>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# **SARS-CoV-2 shedding sources in wastewater and implications for wastewater-based epidemiology**

*Xuan Li<sup>1</sup>, Jagadeeshkumar Kulandaivelu<sup>2</sup>, Ying Guo<sup>1</sup>, Shuxin Zhang<sup>1</sup>, Jiahua Shi<sup>1,3</sup>, Jake O'Brien<sup>4</sup>, Sudipti Arora<sup>5</sup>, Manish Kumar<sup>6</sup>, Samendra P. Sherchan<sup>7</sup>, Ryo Honda<sup>8</sup>, Greg Jackson<sup>4</sup>, Stephen P. Luby<sup>9</sup>, Guangming Jiang<sup>1,3</sup> \**

<sup>1</sup> School of Civil, Mining and Environmental Engineering, University of Wollongong, Australia

<sup>2</sup> Urban utilities, Queensland, Brisbane, Australia

<sup>3</sup> Illawarra Health and Medical Research Institute (IHMRI), University of Wollongong, Wollongong, Australia

<sup>4</sup> Queensland Alliance for Environmental Health Sciences, The University of Queensland, Woollongabba, Queensland 4072, Australia

<sup>5</sup> Dr. B. Lal Institute of Biotechnology, 6E, Malviya Industrial Area, Malviya Nagar, Jaipur -302017, India

<sup>6</sup> Discipline of Earth Science, Indian Institute of Technology Gandhinagar, Gujarat 382 355, India

<sup>7</sup> Department of Environmental health sciences, Tulane University, New Orleans, LA 70112, USA

<sup>8</sup> Faculty of Geosciences and Civil Engineering, Kanazawa University, Kanazawa 920-1192, Japan

<sup>9</sup> Stanford Center for Innovation in Global Health, and Stanford Woods Institute for the Environment, Stanford University, Stanford, California 94305, USA

**\* Corresponding author.**

## **ABSTRACT**

Wastewater-based epidemiology (WBE) approach for COVID-19 surveillance is largely based on the assumption of SARS-CoV-2 RNA shedding into sewers by patients. Recent studies found that SARS-CoV-2 RNA concentration in wastewater ( $C_{\text{RNA}}$ ) could not be accounted by the fecal shedding alone. This study aimed to determine potential major shedding sources based on literature data of  $C_{\text{RNA}}$ , along with the COVID-19 prevalence in the catchment area through a systematic literature review. Theoretical  $C_{\text{RNA}}$  under a certain prevalence was estimated using Monte Carlo simulations, with eight scenarios accommodating feces alone, and both feces and sputum as shedding sources. With feces alone, none of the WBE data was in the confidence interval of theoretical  $C_{\text{RNA}}$  estimated with the mean feces shedding magnitude and probability, and 63% of  $C_{\text{RNA}}$  in WBE reports were higher than the maximum theoretical concentration. With both sputum and feces, 91% of the WBE data were below the simulated maximum  $C_{\text{RNA}}$  in wastewater. The inclusion of sputum as a major shedding source led to more comparable theoretical  $C_{\text{RNA}}$  to the literature WBE data. Sputum discharging behavior of patients also resulted in great fluctuations of  $C_{\text{RNA}}$  under a certain prevalence. Thus, sputum is a potential critical shedding source for COVID-19 WBE surveillance.

## **KEYWORDS:**

COVID-19, SARS-CoV-2, wastewater-based epidemiology, Monte Carlo simulation, virus shedding, sputum.

## **1.INTRODUCTION**

Infection with coronavirus disease (COVID-19) is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The current clinical diagnosis of COVID-19 relies on the testing of individuals by the detection of SARS-CoV-2 RNA using reverse transcription-quantitative polymerase chain reaction (RT-qPCR)(Chau et al., 2020). In contrast to clinical testing, which determines the prevalence by testing individuals, wastewater-based epidemiology (WBE) is a promising approach for population-wide surveillance (Hart and Halden, 2020a; Li et al., 2021a). By systematic collections of wastewater samples either at the inlet of wastewater treatment plants (WWTPs), in sewer pumping stations, or manholes, the WBE approach provides a prevalence estimation with a single wastewater sample for the connected catchment or buildings (Betancourt et al., 2021; Rusiñol et al., 2021; Wong et al., 2021). This requires much lower cost and effort compared with clinical testing for a community with a large population. Depending on the catchment size, the cost of WBE testing accounts for only 0.7-1% of a population-wide clinical testing (Weidhaas et al., 2021). To date, the presence of SARS-CoV-2 RNA in wastewater has been confirmed by numerous studies globally, the meta-analysis of which demonstrated the potential of applying the WBE approach for COVID-19 surveillance (Li et al., 2021b).

Two major approaches have been used for COVID-19 surveillance: 1) conventional prevalence back-estimation based on the concentration of SARS-CoV-2 RNA detected in wastewater samples ( $C_{\text{RNA}}$ ) using WWTPs influent flow rate and inhabitant population (Eq.1) or using the average water consumption data (Eq.2), where feces are regarded as the major source (Ahmed et al., 2020a; Hasan et al., 2020; Saththasivam et al., 2021); 2) infection trend monitoring based on the correlation between the disease incidence (daily new cases, weekly new cases, etc.) and  $C_{\text{RNA}}$  (Graham et al., 2021; Wilder et al., 2021).

$$P_{\text{COVID}} = \frac{C_{\text{RNA}} \times F}{(1-D) \times P \times E} \quad (1)$$

$$P_{\text{COVID}} = \frac{C_{\text{RNA}} \times Q_w}{(1-D) \times E} \times 10^5 \quad (2)$$

Where  $P_{\text{COVID}}$  is the COVID-19 prevalence in the catchment area as the number of COVID-19 patients per 100,000 people;  $C_{\text{RNA}}$  is the concentration of SARS-CoV-2 RNA detected in wastewater samples (gene copies/ L);  $F$  is the daily wastewater flow during the sampling period (L);  $D$  is the decay ratio of SARS-CoV-2 RNA during in-sewer transportation (-);  $P$  is the population size in the catchment area ( $\times 100,000$  people);  $E$  is the excretion rate of SARS-CoV-2 RNA from infected people (gene copies/day·person);  $Q_w$  is the daily water usage that ends up in sewer systems (L/day·person).

Through the conventional back-estimation approach, some studies have estimated the COVID-19 prevalence in the catchment area based on the  $C_{\text{RNA}}$  and fecal shedding parameters including load and probability (Li et al., 2021b; Wölfel et al., 2020). However, in clinical studies, the presence of SARS-CoV-2 RNA in various bodily fluids has been confirmed with feces, urine, blood, saliva, sputum, *etc.* (Table 1). Since SARS-CoV-2 is a respiratory infectious virus, SARS-CoV-2 RNA is mostly detected in respiratory tract samples (typical range 70–100%), to a lesser extent in stool samples (typical range 30–60%), and rarely in urine (<5%) (Jones et al., 2020). Our previous meta-level analysis with more than 1500 patients also revealed that the shedding probability of SARS-CoV-2 RNA in feces and urine was 54.5% and 2.6%, respectively (Li et al., 2021b). The presence of SARS-CoV-2 in stool and urine samples could be related to the swallowing of respiratory secretions from the upper respiratory tract or residues of infected antigen-presenting immune cells, or, more likely, due to virus replication in gastrointestinal epithelial cells or tubular epithelium (Farkash et al., 2020; Moreira et al., 2020; Xiao et al.,

2020b). Thus, considering the shedding probability (Table 1) and the possibility of entering the sewer system, apart from feces, sputum is likely to be an additional major source. A recent study revealed that sputum shedding potentially contributed a great amount of SARS-CoV-2 RNA into wastewater through theoretical simulations (Crank et al., 2022). However, the changes in the prevalence estimation caused by the inclusion of sputum as another shedding source remain unclear.

**Table 1.** Potential shedding sources of SARS-CoV-2 to wastewater

Shedding source	Shedding magnitude (gene copies/g or mL)	Shedding probability	Reference
Feces	$10^{4.52}$ (95% CI: $10^{4.26}$ to $10^{4.78}$ ) copies/g	54.5% (95% CI: 37-73%)	(Li et al., 2021b)
		51.8% (95% CI: 43.8-59.7%)	(van Doorn et al., 2020)
	$10^{3.4}$ (95% CI: $10^{0.24}$ - $10^{6.5}$ ) copies/g		(Miura et al., 2021)
Urine	$10$ - $10^4$ copies/mL	0.026 (95% CI: $6.3 \times 10^{-4}$ -0.10)	(Li et al., 2021b)
		0.8%	(Kim et al., 2020)
Sputum	641 to $1.34 \times 10^{11}$ copies/mL		(Pan et al., 2020)
	$10^{7.3}$ - $10^{8.2}$ copies/mL		(Yoon et al., 2020)

		1/3 of the patients had sputum production	(Lai et al., 2020)
		98% sputum samples tested positive	(Mohammadi et al., 2020)
<b>Saliva</b>	$10^{5.58}$ (95% CI: $10^{5.09}$ to $10^{6.07}$ ) copies/mL	81% (95% CI: 71% to 96%)	(Wyllie et al., 2020)
<b>Blood</b>	8.04 copies/mL	11%	(Peng et al., 2020)

To avoid the uncertainties due to shedding sources, some WBE studies correlated the disease incidence (daily new cases, weekly new cases, etc.) with  $C_{RNA}$  and estimated the current or future incidence in the community based on such correlations (Graham et al., 2021; Wilder et al., 2021). The approach assumes that a constant or similar shedding behavior among the patients. Undoubtedly, nearly all human feces end up in sewers, but the sputum discharging behavior might vary based on the culture, or hygiene practice of patients, especially for patients with upper respiratory infections (a common symptom of COVID-19). To date, the instruction for sputum disposal of patients during self-isolation of COVID-19 is fairly limited. In many countries such as Australia and UK, domestic wastes contaminated with sputum from COVID-19 patients during home isolation are recommended to be double-bagged and stored at home for at least 72 hours or until the clear of COVID-19 symptom before being collected (EPA, 2020; PICA, 2020; Anand et al., 2022). Infectious SARS-CoV-2 was detected in sputum excreted from humans for up to 48 hours in room conditions (Matson et al., 2020), which leads to a potential risk caused by the leakage of sputum if being stored and transported in bags. Thus, due to the inconvenience of waste disposal and long-time storage at home, it is likely that COVID-19

patients are inclined to discharge sputum directly or through handkerchiefs into sewers. However, the impact of personal sputum discharging behavior on the COVID-19 surveillance through WBE approach is largely unknown.

This study aims to investigate the contribution of different shedding sources, primarily feces and sputum, to the SARS-CoV-2 RNA concentration detected in wastewater ( $C_{RNA}$ ) through systematic literature review and Monte Carlo simulation. WBE data of  $C_{RNA}$  with clinically confirmed prevalence in various studies were collected through systematic literature review and compared with the theoretical range based on COVID-19 prevalence. The potential impacts of 1) asymptomatic infected individuals or uncaptured cases by clinical testing; 2) prolonged fecal viruses shedding of recovered cases and 3) recovery efficiency of analytical approach on viral concentration in wastewater were also assessed through simulations. The different simulation scenarios provide a comprehensive evaluation of the major virus shedding sources and their contributions to viral concentrations in wastewater. The knowledge would enhance the understanding of environmental circulation of SARS-CoV-2 in urban water systems and support the application of WBE in the environmental surveillance of COVID-19.

## **2. METHODS**

### **2.1. Systematic literature review of WBE data**

The systematic literature search was conducted on August 4<sup>th</sup>, 2021 following PRISMA guidelines (Silverman and Boehm, 2020). The goal of the search was to collect a comprehensive set of WBE data regarding the  $C_{RNA}$  detected in wastewater and the prevalence of active cases (confirmed cases minus recovered cases) through clinical testing in the catchment area of the wastewater sampling. Databases (i.e., Web of Science core collection, Scopus, and PubMed)

were searched using the term “SARS-CoV-2 AND wastewater AND prevalence”. A total of 602 unique papers were identified after removing duplicates using the EndNote Reference Manager software. Titles and abstracts of the retained articles were screened and assessed for eligibility using these criteria: 1) reported clear data regarding  $C_{RNA}$  and clinically confirmed prevalence in the catchment area; 2) the article is in English and is peer-reviewed. Relevant articles were further assessed by full-text reading and finally, 12 articles with a total of 206 data points were included in this study. Details of the review process are provided in the supplementary information (SI). Other WBE reports were not included in this study due to the lack of required information or inability to provide such information after communicating with their authors.

To date, 14 different RT-qPCR primer-probe sets, targeting various SARS-CoV-2 RNA regions, including nucleocapsid (N), envelope (E), and RNA-dependent RNA polymerase (RdRp), are recommended by WHO and have been applied worldwide (Pezzi et al., 2020; Zhang et al., 2022). In these 12 WBE papers, primer-probe sets targeting at N-gene (i.e. CDC\_N1, CDC\_N2, CDC\_N3; NIID\_N); RdRp-gene (i.e. RdRp\_SARSr, RdRp\_IP4), ORF1ab- gene (ORF1ab) and E-gene (E\_Sarbeco) were applied, which are all in the list of WHO recommendations. However, previous studies found false positives caused by using CDC\_N3, which was then excluded from the US CDC 2019-nCoV RT-qPCR diagnostic panel (CDC, 2020). E\_Sarbeco was not specifically designed for SARS-CoV-2, which would detect other human pathogenic corona viral RNAs such as human coronavirus OC43 (HCoV-OC43), a common cause of mild respiratory tract infection (Park et al., 2021). Thus, considering the specificity, reliability, sensitivity, and WHO recommendations of different primer and probe sets, the results obtained through CDC\_N1, CDC\_N2, RdRp\_SARSr, RdRp\_IP4, and ORF1ab in these WBE papers were included in this study. The performance of these primer-probe sets was thoroughly compared in

previous studies, where the Ct values or  $C_{\text{RNA}}$  detected in the sample were mostly independent of the primer-probe selection through in-laboratory and inter-laboratory comparisons (Fischer et al., 2021; Jung et al., 2020). Thus, the  $C_{\text{RNA}}$  detected in these WBE reports were included without differentiating the primer-probe sets in this study. The clinically testing ratio and testing practice of the country during the WBE surveillance period was collected through the database (<https://ourworldindata.org/coronavirus-testing>) and summarized in Table S1. The impact of sampling technique on the variations of  $C_{\text{RNA}}$  showed conflicting results from the literature. A 10-fold increase in  $C_{\text{RNA}}$  from 24h composite sampling than that of corresponding grab sampling was observed in primary effluent samples, presumably highlighting diurnal variability in the SARS-CoV-2 signal (Gerrity et al., 2021). In contrast, another study found that the sampling technique showed negligible impacts on  $C_{\text{RNA}}$ , where a good agreement between most grab samples and their respective composite samples was observed (Curtis et al., 2021). Thus, to avoid unnecessary loss of data points, studies using either grab sampling or composite sampling were all included in this study. The sampling and analytical methods (including the RT-qPCR primer-probe sets) applied in these 12 articles were summarized in Table S2. The average daily water usage ( $Q_w$ , L/person-day) in each article was collected through the governmental reports for the investigated regions at the year of the study if available; otherwise, the data from the most recent year was included. The wastewater temperature in each study was estimated from (Hart and Halden, 2020b) based on the country and sampling day.

## **2.2. Monte-Carlo simulation of the theoretical $C_{\text{RNA}}$ range**

The Monte Carlo simulation is commonly used when the exact value of results cannot be computed with deterministic algorithms. The principle behind the Monte Carlo methodology is the law of large numbers in probabilistic statistical theory, where the frequency of the random

event is approximately equal to the probability of event occurrence after repeated trials (Zhao et al., 2017). Therefore, by taking a prescribed number of samples from defined distributions for model input parameters, a Monte Carlo simulation provides probabilities of different outcomes occurring in an estimated probability distribution (Gilks, 2005).

Our previous study found that estimating the prevalence using Eq. 2 showed lower uncertainty in comparison to Eq.1, due to the lower uncertainty of average water consumption data (used in Eq.2) than that of WWTPs influent flow rate and inhabitant population (used in Eq.1) (Li et al., 2021b). Thus, Eq.2 was applied to simulate the theoretical  $C_{RNA}$  range for a certain COVID-19 prevalence. Recent studies reported that SARS-CoV-2 RNA decay in wastewater followed the first-order kinetics (Eq.3), where the time in wastewater and decay rate constant ( $k$ , increasing with higher temperature) were critical factors (Ahmed, 2020; Bivins et al., 2020). Depending on the wastewater temperature, the  $k$  values ranged from 0.067-0.286/day under different temperatures (Ahmed, 2020; Bivins et al., 2020). In this study, to simplify the simulation, the  $k$  value observed at 20-25 °C (common wastewater temperature of the WBE studies included in this paper as described in section 2.2) was applied as 0.1/day. Thus, Eq. 4 and Eq.5 were subsequently established for the scenario with feces as the only shedding source, and both feces and sputum as the major shedding sources, respectively.

$$C_{RNA,t} = C_{RNA,0} \times e^{-kt} \quad (3)$$

$$C_{RNA} = \frac{P_{COVID} \times e^{-0.1 \times HRT} \times P_s \times Q_s \times 10^{Rs}}{Q_w} \quad (4)$$

$$C_{RNA} = \frac{P_{COVID} \times e^{-0.1 \times HRT} \times (P_s \times Q_s \times 10^{Rs} + P_{sp} \times Q_{sp} \times 10^{Rsp} \times C_f)}{Q_w} \quad (5)$$

Where  $C_{RNA,t}$  and  $C_{RNA,0}$  are the concentrations of SARS-CoV-2 RNA (gene copies/L) in wastewater at time  $t$  and time 0, respectively, and  $k$  is the decay rate constant (1/day) (Ahmed,

2020) (Ahmed, 2020) (Ahmed, 2020) (Ahmed, 2020).  $P_s$  is the shedding probability in feces from a COVID-19 patient (-);  $Q_s$  is the daily amount of feces of an individual (g/person·day);  $R_s$  is the logarithmic shedding magnitude of SARS-CoV-2 RNA in feces ( $\log_{10}$ , gene copies/g), HRT (hydraulic retention time) is the in-sewer transportation time (day).  $P_{sp}$  is the shedding probability in sputum samples from a COVID-19 patient (-);  $Q_{sp}$  is the daily shedding amount of sputum of an individual (mL/person·day);  $R_{sp}$  is the logarithmic shedding magnitude of SARS-CoV-2 RNA in sputum samples ( $\log_{10}$ , gene copies/mL),  $C_f$  is the ratio of sputum that enters sewers (-).

For WBE surveillance of COVID-19, the population-wide viral RNA shedding information is critical. SARS-CoV-2 RNA shedding magnitude, probability for each shedding source among patients are largely impacted by physiological factors such as gender, age, and pathological conditions (Novazzi et al., 2020; Wang et al., 2020). Our previous meta-analysis summarized the clinical results from around 1500 patients covering all the gender, age groups (children, adults<60 and adults over 60), and pathological conditions (severe, moderate, and mild), which revealed that the mean shedding magnitude was  $10^{4.523 \pm 0.133}$  gene copies/g, and the mean shedding probability ( $P_s$ ) was  $0.545 \pm 0.093$  (Li et al., 2021b). Similar values were reported in other meta-level analyses where the positive proportion of the fecal samples was found to be 51.8% (95% CI 43.8-59.7%) and the median shedding concentration was  $10^{3.4}$  (95%CI:  $10^{0.24}$ - $10^{6.5}$ ) to  $10^{4.9}$  (95%CI:  $10^{3.9}$ - $10^{6.8}$ ) (Crank et al., 2022; Miura et al., 2021; van Doorn et al., 2020). The detection efficiency of clinical protocol for SARS-CoV-2 RNA in stool samples was confirmed using standardized stool samples (stool samples with no SARS-CoV-2 RNA) spiked with synthetic SARS-CoV-2 RNA, where comparable results to the initial spike-in concentrations were achieved (Poon and Tee, 2021), suggesting that the detected SARS-CoV-2

RNA concentration can largely reflect the actual concentration in stool samples. Thus, the distribution for  $R_s$  and  $P_s$  was applied as Gaussian distributions in the form of Normal ( $\mu, \sigma^2$ ): i.e., Normal (4.523, 0.017) and Normal (0.545, 0.009), respectively. The probability density of  $R_s$  and  $P_s$  during the simulations was detailed in Figure S1 and S2. The median feces mass was around 150-250 g/person·day in healthy individuals (Rose et al., 2015). Gastrointestinal conditions have been observed in 12% of COVID-19 patients although the feces amount generated has not been reported (Walsh et al., 2020). Thus, feces mass was applied as 200 and 300 g/person·day (to accommodate the presence of fecal shedding of COVID-19 patients with gastrointestinal conditions in the community) in this study to simulate the minimum and maximum shedding conditions.

The contribution of daily water usage into wastewater flow varied regionally, based on the WBE data collected (described in section 2.1), the range was applied as 150-400 L/day·person. In-sewer transportation time, or more specifically hydraulic retention time (HRT) of sewers, could also affect the decay of SARS-CoV-2 (Eq.3). Due to the lack of or inability to obtain HRT information of WBE studies (described in section 2.1) and the unknown distribution pattern of COVID-19 patients in the catchment area, two HRTs, i.e. 30 minutes and 10 hours, were applied to reflect the conditions of a small and large scale of WWTPs (McCall et al., 2017). Thus, two scenarios were applied for the theoretical  $C_{RNA}$  simulation with feces shedding: F1 (lower boundary) with the shortest HRT, the highest  $Q_s$ , and the lowest  $Q_w$ , and F2 (upper boundary) with longest HRT, the lowest  $Q_s$ , and the highest  $Q_w$  (Table 2). The maximum SARS-CoV-2 RNA concentration in clinical stool samples of COVID-19 patients ranged from  $10^8$ - $10^{8.5}$  copies/g in clinical reports and meta-analysis (Jones et al., 2020; Wölfel et al., 2020). Thus, the maximum  $C_{RNA}$  with feces as the only shedding source was simulated with the maximum

concentration of SARS-CoV-2 RNA detected clinically in feces samples ( $10^{8.5}$  copies/g), the shortest HRT, the highest  $Q_s$ , and the lowest  $Q_w$  in scenario FM (Table 2).

**Table 2.** The range of input parameters of  $C_{RNA}$  simulation for seven scenarios

Shedding source	Simulation scenario		Code	Parameters
Feces	Mean shedding magnitude and probability	Upper boundary	F1	HRT=30 minutes, $Q_s=300$ g, $Q_w=150$ L, $P_s \sim \text{Normal}(0.545, 0.009)^a$ , $R_s \sim \text{Normal}(4.523, 0.017)$
		Lower boundary	F2	HRT=10 h, $Q_s=200$ g, $Q_w=400$ L, $P_s \sim \text{Normal}(0.545, 0.009)$ , $R_s \sim \text{Normal}(4.523, 0.017)$
	Maximum shedding magnitude	Maximum $C_{RNA}$ concentration	FM	HRT=30 minutes, $Q_s=300$ g, $Q_w=150$ L, $P_s \sim \text{Normal}(0.545, 0.009)$ , $R_s=8.5$
Sputum and feces	Sputum shedding magnitude follows Normal (6, 0.6)	Lower boundary (20% sputum entering sewers)	SPS1-1	HRT=6 h, $Q_s=300$ g, $Q_w=150$ L, $P_s \sim \text{Normal}(0.545, 0.009)$ , $R_s \sim \text{Normal}(4.523, 0.017)$ , $P_{sp} = 0.33$ , $R_{sp} \sim \text{Normal}(6, 0.6)$ , $Q_{sp} \sim \text{Normal}(95, 9.5)$ , $C_f=0.2$
		Upper boundary (80% sputum entering sewers)	SPS1-2	HRT=6 h, $Q_s=300$ g, $Q_w=150$ L, $P_s \sim \text{Normal}(0.545, 0.009)$ , $R_s \sim \text{Normal}(4.523, 0.017)$ , $P_{sp} = 0.33$ , $R_{sp} \sim \text{Normal}(6, 0.6)$ , $Q_{sp} \sim \text{Normal}(95, 9.5)$ , $C_f=0.8$
	Sputum shedding magnitude follows Normal (8, 0.8)	Lower boundary (20% sputum entering sewers)	SPS2-1	HRT=6 h, $Q_s=300$ g, $Q_w=150$ L, $P_s \sim \text{Normal}(0.545, 0.009)$ , $R_s \sim \text{Normal}(4.523, 0.017)$ , $P_{sp} = 0.33$ , $R_{sp} \sim \text{Normal}(8, 0.8)$ , $Q_{sp} \sim \text{Normal}(95, 9.5)$ , $C_f=0.2$
		Upper boundary (80% sputum entering sewers)	SPS2-2	HRT=6 h, $Q_s=300$ g, $Q_w=150$ L, $P_s \sim \text{Normal}(0.545, 0.009)$ , $R_s \sim$

entering sewers)			Normal (4.523, 0.017), $P_{sp} = 0.33$ , $R_{sp} \sim \text{Normal}(8, 0.8)$ , $Q_{sp} \sim \text{Normal}(95, 9.5)$ , $C_f = 0.8$
Maximum shedding magnitude and amount	Maximum $C_{RNA}$ concentration	SPSM	HRT=6 h, $Q_s=300$ g, $Q_w=150$ L, $P_s \sim \text{Normal}(0.545, 0.009)$ , $R_s=8.5$ , $P_{sp} = 0.33$ , $R_{sp}=11.1$ , $Q_{sp} \sim \text{Normal}(95, 9.5)$ , $C_f=0.8$

Note: <sup>a</sup>: Normal distribution is in the form of Normal ( $\mu, \sigma^2$ ). The parameter  $\mu$  is the mean or expectation of the distribution, while the parameter  $\sigma$  is its standard deviation.

To date, the understanding of SARS-CoV-2 RNA shedding in sputum samples is limited. A study assessed the shedding magnitude in sputum samples of 80 patients, where the viral loads varied from 641 to  $1.34 \times 10^{11}$  gene copies/mL with a median value of  $7.52 \times 10^5$  gene copies/mL (Pan et al., 2020). A similar range of  $10^{7.3}$ - $10^{8.2}$  gene copies/mL was observed in another study with two patients (Yoon et al., 2020). To reflect the population-wide shedding, the  $R_{sp}$  (logarithmic shedding magnitude of SARS-CoV-2 RNA in sputum) were assumed as normal distributions in the form of ( $\mu, \sigma^2$ ). In these reports, the  $\sigma^2$  of  $R_{sp}$  was observed as 5.4% of  $\mu$  ( $\mu=7.78$ ) (Yoon et al., 2020) and 9.3% of  $\mu$  ( $\mu=5.88$ ) (Pan et al., 2020). Thus, two shedding magnitudes were assumed as normal distributions in the form of ( $\mu, \sigma^2$ ) with 10% of  $\mu$  applied for  $\sigma^2$ , as Normal (6, 0.6) for SPS1-1 and SPS1-2, and Normal (8, 0.8) for SPS2-1 and SPS2-2, respectively (Table 2). SPS1-1, SPS1-2, SPS2-1, SPS2-2 represent the simulations when sputum is considered as a major shedding source in addition to feces.

A recent meta-analysis of 970 patients revealed that sputum production was observed in one third of the patients (Lai et al., 2020) and the SARS-CoV-2 positive ratio was 98% in sputum samples of patients at 0-7 days after the symptom onset (Mohammadi et al., 2020). Based on the limited data,  $P_{sp}$  was determined as 0.33 ( $98\% \times 1/3$ ) to reflect the shedding probability among

patients. Healthy individuals generally do not produce a large amount of sputum (Balbi et al., 2007) but the daily amount of sputum produced in patients with respiratory diseases was  $94.6 \pm 21.6$  mL (Lin et al., 1997). Currently, the amount of daily sputum production in SARS-CoV-2 patients has not been reported yet. Considering the respiratory symptom of COVID-19 patients,  $Q_{sp}$  was adopted as Normal (95, 9.5) (Table 2). The amount of sputum entering sewers has not been investigated to date. Due to the complexity of waste disposal as recommended by the health departments, it is more likely a large portion of sputum from COVID-19 patients is discharged into sewers. To explore the lower and higher possibilities, the percentage of sputum entering sewers ( $C_f$ ) was set as 0.2 (SPS1-1, SPS2-1) and 0.8 (SPS1-2, SPS2-2), respectively. To simplify the input scenarios, the common HRT in sewers (i.e., 6 h), the highest  $Q_s$ , and the lowest  $Q_w$  were applied for the scenarios using both sputum and feces as virus shedding sources into wastewater (Table 2).

Using Eq. 4 and Eq. 5, the theoretical  $C_{RNA}$  range with 8 input scenarios as described in Table 2 was simulated through Monte Carlo models using OpenBUGS (version 3.2.3). The details of the models are included in the SI. To provide stable distributions of results, an initial simulation of 5,000 iterations was first performed as a burn-in, and a further 5,000 iterations were used for calculating outputs of the model simulations.

### **2.3. Comparison between detected $C_{RNA}$ from WBE studies and simulated $C_{RNA}$**

The  $C_{RNA}$  detected in wastewater with corresponding prevalence (clinically confirmed active COVID-19 cases per 100,000 people in the catchment area) was compared with the theoretical range simulated in different scenarios as described in section 2.2. The number of data fitted into each scenario was counted. The difference between detected  $C_{RNA}$  and simulated  $C_{RNA}$  with the same prevalence in each scenario was evaluated by the root mean squared error (RMSE).

## 2.4 Assess the contribution of sputum to the total SARS-CoV-2 RNA shedding in wastewater

To further assess the contribution of sputum to the overall SARS-CoV-2 RNA shedding (feces and sputum together), the contribution of sputum shedding ( $CR$ ) was calculated as Eq.6.

$$CR = \frac{P_{sp} \times Q_{sp} \times 10^{R_{sp}} \times C_f}{P_s \times Q_s \times 10^{R_s} + P_{sp} \times Q_{sp} \times 10^{R_{sp}} \times C_f} \quad (6)$$

For fecal shedding, the maximum shedding quantity ( $Q_s$ , 300g), the mean logarithmic shedding magnitude ( $R_s \sim \text{Normal}(4.523, 0.017)$ ), and the mean shedding probability ( $P_s \sim \text{Normal}(0.545, 0.009)$ ) were applied. For sputum shedding, the shedding probability ( $P_{sp}$ ) and shedding quantity ( $Q_{sp}$ ) was applied as 0.33 and  $\text{Normal}(95, 9.5)$ , respectively. Seven different ratios of sputum that enters sewers ( $C_f$ ) (i.e., 0.01, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8) and two shedding magnitude ( $R_{sp} \sim \text{Normal}(6, 0.6)$ ,  $R_{sp} \sim \text{Normal}(8, 0.8)$ ) were used to reflect the different shedding scenarios (14 scenarios in total). The mean of  $CR$  with standard errors under each shedding scenario was simulated through Monte Carlo models using OpenBUGS (version 3.2.3). The details of the models are included in the SI. To provide stable distributions of results, an initial simulation of 5,000 iterations was first performed as a burn-in, and a further 5,000 iterations were used for calculating outputs of the model simulations.

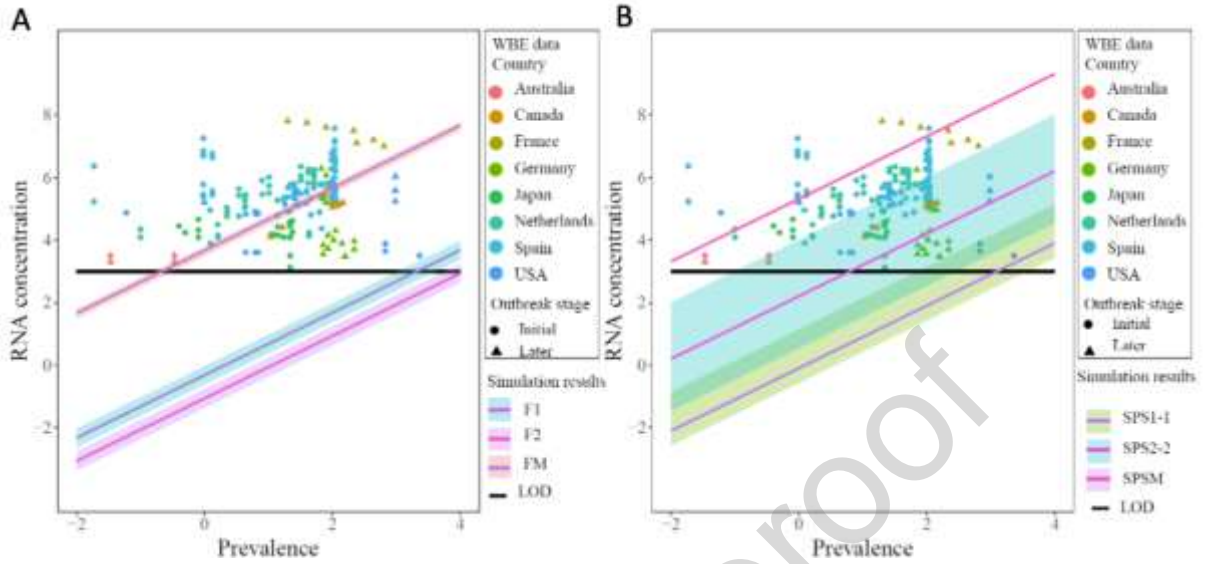
## 3. RESULTS

### 3.1 Comparison between detected $C_{RNA}$ from WBE studies and simulated $C_{RNA}$ with and without sputum included as a shedding source

During the time span for each study, although these countries had different contacting tracing and clinical testing policies (text S1.1), the testing positive rates (daily positive cases/daily total tests) were all lower than 10% (Table S1), indicating a good coverage of the clinical testing

(Saglietto et al., 2020). The WBE data were collected from eight different countries with the outbreak at the initial stage (within 1 detection window (i.e. 28 days) of SARS-CoV-2 for wastewater samples after the first COVID-19 patient) or later stage (after the initial stage) (Figure 1). At the initial stage, the number of active cases equals cumulative cases in the catchment area. At the later stage, the prevalence of COVID-19 in the catchment area was calculated based on the active cases (confirmed cases minus recovered cases).

With feces as the only shedding source, the theoretical  $C_{RNA}$  range was estimated based on the mean shedding magnitude and probability in F1 (upper boundary) and F2 (lower boundary) (Figure 1A). None of the WBE data points was in the range of F1 and F2, suggesting that the  $C_{RNA}$  detected were all higher than the theoretically estimated values. Furthermore, 63% of the reported WBE data points were above the confidence interval of FM (Figure 1A). The data above the confidence interval of FM did not show a clear correlation with the outbreak stage or the country (Figure 1A), suggesting consistent higher shedding loads than fecal shedding, regardless of the regional or temporal difference. FM was established based on the highest fecal shedding magnitude, the shortest HRT, maximum daily feces mass amount and the lowest daily water usage (Table 2). The feces shedding magnitude in FM was only observed in some clinical cases during a short period (2 days) as reported by (van Doorn et al., 2020). In addition, the limit of detection (LOD) of SARS-CoV-2 RNA was observed to be around  $10^3$  copies/L (Ahmed et al., 2020b). With the mean shedding magnitude, only COVID-19 prevalence higher than 1000/100,000 can be captured by WBE (F1 and F2 in Figure 1A), which conflicts with the successful detection of  $C_{RNA}$  in wastewater under much lower prevalence.



**Figure 1.** Comparison of reported WBE data against the simulated  $C_{RNA}$  with 95% confidence interval under a certain prevalence for feces as the shedding source (A) and both feces and sputum as shedding sources (B). The  $C_{RNA}$  (copies/L) and prevalence (clinically confirmed COVID-19 cases/100,000 person) are illustrated on logarithmic scales. The color and shape of the points represent the country, and the outbreak stage of the WBE data, respectively. The initial stage of the outbreak is within 1 detection window (i.e. 28 days) of SARS-CoV-2 for wastewater samples after the first COVID-19 patient, and the later stage is after 1 detection window of the first patients in the catchment area (Ahmed et al., 2020a).

The maximum  $C_{RNA}$  with both feces and sputum as the shedding source was estimated with the highest sputum ( $10^{11.1}$  copies/mL) and fecal shedding magnitude ( $10^{8.5}$  copies/g) observed in clinical cases (SPSM in Figure 1B). The fixed shedding magnitude rather in SPSM than a Gaussian distribution in other scenarios (SPS1-1 to SPS 2-2) leads to a narrow confidence interval in SPSM. The region below SPSM reflects the greatest probability of the occurrence of WBE data. 91% of the WBE data fell into the range below SPSM, in contrast to that of 37% in FM (Figure 1A). This suggests that sputum was likely another critical shedding source for

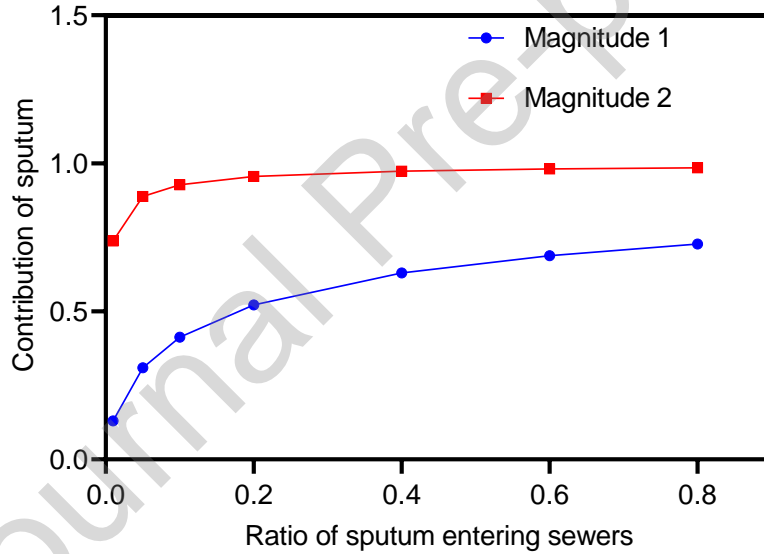
SARS-CoV-2 into sewers, and a high proportion of sputum generated by COVID-19 patients was potentially discharged into sewers. Four scenarios were applied for estimating the  $C_{RNA}$  under a certain prevalence with both sputum and feces as major shedding sources (Table 2). For better readability, the highest (SPS2-2) and lowest (SPS1-1) theoretical  $C_{RNA}$  range were included in Figure 1B, and two other scenarios (SPS1-2, and SPS2-1) are included in Figure S3.

SPS1-1 and SPS1-2, and SPS2-1 and SPS2-2 were simulated under the sputum shedding magnitude of Normal (6, 0.6) and Normal (8, 0.8), respectively, where the upper (SPS1-2, SPS2-2) and lower boundary (SPS 1-1, SPS 2-1) under each shedding magnitude were estimated by assuming 20% and 80% of sputum entering sewers, respectively (Table 2). In these four scenarios, only three of the WBE data were in the confidence interval of SPS1-1, while about 49% of the WBE data was in the range of SPS2-2 (Figure 1B). The RMSE of logarithmic  $C_{RNA}$  was 3.73, 3.20, 1.84, and 1.45 in SPS1-1, SPS1-2, SPS2-1, and SPS2-2, respectively. Thus, the range of  $C_{RNA}$  estimated under a higher sputum shedding magnitude and amount entering sewers was more comparable to the concentration detected in WBE studies. More importantly, when the sputum shedding load (magnitude  $\times$  amount entering sewers) increased by about 100 times, the median  $C_{RNA}$  increased by about 50 times from SPS1-1 to SPS 2-1, but about 100 times from SPS1-2 to SPS 2-2 (Figure 1B, Figure S3). This implies a dominant role of sputum shedding at higher shedding loads for the prevalence estimation through WBE approach.

### **3.2 Contribution of sputum to the total SARS-CoV-2 RNA shedding in wastewater**

The contribution of sputum discharging behaviors of patients to the overall SARS-CoV-2 RNA shedding (feces and sputum together) was further evaluated under different ratios of sputum entering sewers with two shedding magnitudes (Magnitude1 and Magnitude 2) (Figure 2). It is evident that sputum shedding played a critical role in the total virus shedding, especially under

higher shedding magnitude and or higher ratio entering the sewers. Under the lower shedding magnitude (Magnitude 1), sputum may contribute from 13% to 73% of the total virus in wastewater, when the ratio of sputum entering sewers increase from 1% to 80%. With the higher shedding magnitude of sputum (Magnitude 2), even with 1% of sputum entering the sewers, the contribution of sputum reached 74% of the total virus RNA in wastewater. Furthermore, sputum became the dominant shedding source (>90% contribution) with >10% of sputum entering sewers under the higher shedding load (Magnitude 2), where the fecal shedding became negligible.



**Figure 2.** Contribution of sputum to the total virus shedding under two sputum shedding magnitudes. Magnitude 1 and Magnitude 2 represent  $R_{sp} \sim \text{Normal}(6, 0.6)$ , and  $R_{sp} \sim \text{Normal}(8, 0.8)$ , respectively. The error bar represents the standard error of 5000 simulation results. For some data points, the error bars are shorter than the height of the symbol and thus are not visible. The personal sputum discharging practice also greatly affected the range of theoretical  $C_{RNA}$ . With both feces and sputum as shedding sources, the increase of sputum entering sewer from 20% to 80% led to 2.5 times higher median  $C_{RNA}$  with the lowest shedding magnitude, but 4 times

increase in  $C_{RNA}$  with the highest shedding magnitude (Figure S3). More importantly, with the least shedding magnitude of sputum (Magnitude 1), 20% of sputum generated entering sewers would increase the  $C_{RNA}$  by about 2 times compared with only fecal shedding. With a higher shedding magnitude (Magnitude 2), even 1% of sputum entering sewers would increase the  $C_{RNA}$  by 5 times. Higher sputum shedding loads (magnitude  $\times$  amount entering sewers) would lead to up to 70 times increase of the  $C_{RNA}$  in comparison to the results from fecal shedding alone.

## **4. Discussion**

### **4.1. Feces contributes partly to detected SARS-CoV-2 RNA in wastewater**

Through Monte-Carlo simulations, the  $C_{RNA}$  detected in WBE studies showed much higher values than their theoretical range under corresponding prevalence. More importantly, 63% of the reported WBE data points were above the confidence interval of the maximum shedding condition (FM). Previous studies also noticed that the  $C_{RNA}$  detected in wastewater was higher than the theoretical values estimated by the clinically confirmed prevalence using feces as the shedding source (Wu et al., 2020a; Wu et al., 2020b). This was attributed to a hypothetical surge of shedding (either from feces or other sources) before the symptom onset at several orders of magnitude greater than typical values. However, a recent study monitored the  $C_{RNA}$  in wastewater of a university campus with intense on-campus case surveillance, which revealed that individual shedding of RNA (sources unclear) into wastewater peaks on average six days after the symptom onset (95% UI: 4 – 8 days) (Cavany et al., 2021). Also, most of the virological assessments were carried out on patients with confirmed symptoms or clinical tests by oropharyngeal or nasopharyngeal swabs. Thus, the hypothetical surge of shedding before symptom onset is debatable.

Some other factors, such as the shedding from asymptomatic patients or uncaptured cases by clinical testing, analytical recovery efficiency, in-sewer RNA decay, and duration shedding could also affect the  $C_{\text{RNA}}$  detected in the wastewater although they were not included in the simulation. Thus, the differences between simulation results and the actual  $C_{\text{RNA}}$  detected in the wastewater, and the contributions of the above factors were further compared and discussed (Table 3).

**Table 3.** Potential contributions of additional factors to the simulated wastewater concentration of SARS-CoV-2 RNA

Factors	Value	Changes to the simulated $C_{\text{RNA}}$	
		Increase (times)	Decrease (times)
Percentage of asymptomatic patients or uncaptured cases	15.6%-76.7%	0.2-3.3	-
Analytical recovery efficiency	10%-72%	-	0.4-9.0
In-sewer RNA decay	$k=0.084\text{-}0.286/\text{day}$ , HRT=1 minute to 10 hours	0-1.0	0-1.0
Shedding duration	Up to four to ten weeks after the first symptom onset	-	-
Total		0.2-7.6	0.4-19.0

SARS-CoV-2 virus shedding has been found in feces of asymptomatic patients (Park et al., 2020; Tang et al., 2020; Zhang et al., 2020a; Zhang et al., 2020b). The COVID-19 prevalence was predominately confirmed through clinical testing of individuals. Symptom-onset can be a major trigger for the motivation of testing, in addition to some mandatory testing required for cross-border travelers or close contacts of infected patients (Table S1). A meta-analysis conducted in July 2020 (similar period as the WBE studies) estimated that the percentage of asymptomatic

patients among COVID-19 patients was 15.6% (95% CI, 10.1%-23.0%) (He et al., 2021). The Centre for Disease Control and Prevention (CDC) also estimated that 1 in 4.3 (95% UI 3.7-5.0) of total COVID-19 infections were reported (Reese et al., 2020). Therefore, considering the potential shedding from asymptomatic patients or uncaptured cases, the theoretical  $C_{RNA}$  would increase by 0.2-3.3 times.

The analytical approach (i.e. concentration, extraction, and detection) applied for wastewater could greatly affect the recovery efficiency of the SARS-CoV-2 RNA from wastewater, which thereby changes the  $C_{RNA}$  detected in wastewater (Li et al., 2021b; Rusiñol et al., 2020). The large variation of  $C_{RNA}$  is partly due to the limitations of wastewater analysis using RT-qPCR, as shown by the SARS-CoV-2 interlaboratory consortium report (Pecson et al., 2021). Generally, to further correct the  $C_{RNA}$  detected in wastewater, the recovery efficiency is quantified by spiking low-pathogenic surrogate viruses as external controls or using internal controls such as fecal load indicators (e.g. pepper mild mottle virus) (Ahmed et al., 2020c; Jafferli et al., 2020; Wu et al., 2020b). The recovery efficiency of the included articles varied from 10% to 72% (Table S2). Thus, the analytical approach could lead to 0.4-9 times decrease of theoretical  $C_{RNA}$  (Table 3).

The decay of SARS-CoV-2 RNA during in-sewer transportation was found to follow the first-order kinetics as Eq.3, with  $k$  values increased from 0.084/day to 0.286/day from 4°C to 37°C (Ahmed, 2020). Apart from the  $k$  value, the traveling time  $t$  (HRT of sewers) also impacts the decay of SARS-CoV-2 RNA in wastewater (Eq. 3). For most of the WWTPs, the HRT of sewers ranged from several minutes to 6-10 hours in small and large scale WWTPs, respectively (McCall et al., 2017). In this study, the  $k$  (0.1/day) value at 20°C (common wastewater temperature in these WBE studies) was applied for simulation. The HRT of 30 minutes and 10 hours was considered to simulate the upper (F1) and lower boundary (F2) for feces as the only

shedding source. Thus, changes in wastewater temperature or HRT would potentially lead to a further 1-time increase for the upper boundary ( $k=0.286/\text{day}$ ,  $\text{HRT}=1\text{-}30$  minutes), or a 1-time decrease for the lower boundary ( $k=0.084/\text{day}$ ,  $\text{HRT}=10$  hours) (Table 3). In addition, prolonged fecal shedding has been observed in patients for up to four to ten weeks after the first symptom onset and even after the patients' respiratory samples tested negative for SARS-CoV-2 RNA (Wu et al., 2020c; Xiao et al., 2020a; Xu et al., 2020; Yang et al., 2020). However, a meta-analysis involving thousands of patients revealed that the shedding loads of recovered patients (after respiratory samples tested negative) was about 0.3%-6.2% of the active patients (Jones et al., 2020). Thus, the contribution of prolonged shedding of recovered cases is negligible in comparison to active cases.

Compared with the reduction of  $C_{\text{RNA}}$  due to the analytical recovery efficiency, the increase of  $C_{\text{RNA}}$  caused by asymptomatic patients or uncaptured cases, and in-sewer decay is limited (Table 3). Even with 100% analytical recovery, the inclusion of all the factors would elevate the simulation results of  $C_{\text{RNA}}$  by 0.2-7.6 times (Table 3). With elevated maximum shedding condition ( $\text{FM}\times 7.6$ ), 40% of the reported WBE data points were still above the confidence interval. Furthermore, the  $C_{\text{RNA}}$  detected in WBE studies was averagely  $10^{2.81}$  to  $10^{4.20}$  times higher than the upper (F1) and lower (F2) boundary of fecal shedding, respectively. Some  $C_{\text{RNA}}$  detected in wastewater were even about  $10^4$  times higher than the theoretical maximum  $C_{\text{RNA}}$  with feces as the only shedding source (Figure 1A). Thus, feces only contribute partly as a shedding source to SARS-CoV-2 RNA in wastewater.

#### **4.2. Potential role of sputum as a major SARS-CoV-2 RNA shedding source**

Through the simulation, it is evident that the range of  $C_{\text{RNA}}$  estimated under a higher sputum shedding magnitude was more comparable to the concentration detected in WBE studies. At

higher shedding loads (magnitude  $\times$  amount entering sewers), sputum became the dominant source, where the impact of fecal shedding becomes negligible. Thus, the hygiene behavior of patients in regards to the disposal of sputum becomes more important for the WBE surveillance of COVID-19.

As aforementioned, a clear shedding load or constant shedding behavior from patients is crucial, either for the WBE surveillance of COVID-19 through the conventional back-estimation approach or the correlation between  $C_{RNA}$  and COVID-19 prevalence or incidence. Our simulation results suggest that, with a higher shedding magnitude, even 1% of sputum entering sewers, the  $C_{RNA}$  would be increased by 5 times. Higher sputum shedding loads (magnitude  $\times$  amount entering sewers) would lead to up to 70 times increase of the  $C_{RNA}$  in comparison to the results from fecal shedding alone. Even with both feces and sputum considered as shedding sources, the increase of sputum entering sewers from 20% to 80% led to 2.5 times higher median  $C_{RNA}$  with the lowest shedding magnitude, but 4 times increase in  $C_{RNA}$  with the highest shedding magnitude. Thus, the change of person sputum discharge behavior would introduce great variations into the WBE surveillance of COVID-19. Highly over-dispersed viral shedding was also observed in some WBE studies (Cavany et al., 2021), which is likely related to the personal sputum discharging behaviors. To date, the instructions regarding sputum disposal remain unclear yet for COVID-19 patients. Although discharging sputum into sewers is intuitively safer during self-isolation, the actual amount of sputum entering sewers is likely dispersed and varied greatly among the population. Thus, future research regarding the sputum discharging behavior among patients and clear instructions for sputum disposal are crucial for the WBE surveillance of COVID-19.

In addition, with both feces and sputum as the major shedding source, 9% of WBE data points were above SPSM (maximum fecal+sputum shedding) (Figure 1B). The RMSE of logarithmic  $C_{RNA}$  with both sputum and feces as the shedding source ranged from 1.45-3.73 with the change of shedding loads of sputum (SPS1-1 to SPS2-2). As discussed in section 4.1, other factors such as the shedding from asymptomatic patients or uncaptured cases by clinical testing, and in-sewer RNA decay as listed in Table 3 would also increase the  $C_{RNA}$  detected in the wastewater by up to 7.6 times. This could further reduce the RMSE of logarithmic  $C_{RNA}$  to 1.25-2.88, which is more comparable to the WBE data, resulting in 96% of WBE data falling below the elevated SPSM ( $SPSM \times 7.6$ ). The remaining 4% of WBE data that was still above the elevated SPSM might be caused by the presence of other shedding sources. SARS-CoV-2 RNA was also detected in the water used for mouth/throat wash of COVID-19 patients at around  $10^2$ - $10^5$  gene copies/mL (Liu et al., 2020). This could be caused by the presence of SARS-CoV-2 RNA in saliva, which might be another potential source of SARS-CoV-2 RNA in wastewater (Huang et al., 2021). The SARS-CoV-2 RNA load and detection rate in saliva were found comparable to that of sputum and respiratory tract samples (Zhu et al., 2020). However, due to the limited understanding of the concentration, detection probability, and daily discharge amount of mouth/throat wash or saliva into sewers, the contribution of mouth/throat wash and saliva was not included in the simulation, which requires future investigations.

The country and outbreak stage showed a negligible difference in the distribution of WBE data in comparison to the simulation results from fecal shedding (Figure 1A). However, the WBE studies from the USA and France had more data points above the SPSM (maximum fecal+sputum shedding) (Figure 1B). This might be related to the higher shedding load of SARS-CoV-2 RNA (from either feces, or sputum, or both) or sputum discharging practice from

COVID-19 patients in these countries. To date, the impact of race, socioeconomic conditions, and country on shedding dynamics in COVID-19 patients remains unclear, which requires future investigations.

### **5. Implications, limitations, and future research recommendations**

Our results suggest that feces is unlikely to be the only shedding source based on currently available WBE data. Sputum might play an important role as a shedding source to the sewer system. The theoretical ranges of SARS-CoV-2 RNA in wastewater estimated under higher sputum shedding magnitude (around  $10^8$  copies/mL) were more comparable to the WBE data. Sputum shedding became the dominant source for WBE estimations under such scenarios. The discharge of sputum into sewers can lead to up to 70 times increase of SARS-CoV-2 RNA in wastewater while other factors such as asymptomatic or uncaptured cases by clinically testing, analytical recovery efficiency, in-sewer decay and prolonged shedding from recovered patients have comparably limited impacts on the change of  $C_{RNA}$  in wastewater in comparison to sputum shedding. However, sputum was not previously considered as a major source of SARS-CoV-2 shedding. WBE investigations were primarily focusing on feces detection and quantifications. Due to the lack of such information, the sputum shedding magnitude in this study was based on the data from two studies with limited numbers of patients. The data points of sputum shedding magnitude were comparatively fewer than that of fecal shedding (i.e. 1500 patients). The daily sputum amount generated by a COVID-19 patient was estimated based on patients with respiratory diseases. Furthermore, the amount of sputum entering the sewers will vary depending on the personal hygiene practices and recommendations provided by the relevant health agency. This study provided two scenarios with 20% and 80% of sputum entering sewers and found the percentage of sputum entering sewers became more important under higher sputum shedding

loads. The contribution of sputum was further assessed with 1% to 80% entering sewers. Even 1% of sputum entering sewers under the higher shedding magnitude would increase the  $C_{RNA}$  in wastewater by 5 times, which would greatly impact WBE surveillance of COVID-19. Based on the currently available data on sputum shedding, the simulation results in this study suggest a critical role of sputum shedding on the COVID-19 WBE estimations. A comprehensive survey for the sputum discharge practice of COVID-19 patients and the amount of virus load in the sputum are recommended for future investigations. When such information becomes available, the contribution of sputum shedding would be further validated or evaluated based on stronger clinical evidence.

In addition, to date, there is a lack of assessment regarding the geological and or temporal differences in the contribution of daily water usage into wastewater flow. Since the WBE data included in this study was collected from eight countries (Australia, Canada, France, Germany, Japan, Netherlands, Spain, and the USA), the  $C_{RNA}$  was simulated based on their water usage. In other countries with different water usage cultures, the theoretical  $C_{RNA}$  is likely to be different and can be biased due to limited sanitation coverage.

#### SUPPLEMENTARY INFORMATION

Systematic literature review process for collecting WBE data. COVID-19 testing coverage in selected countries. Sampling and analytical techniques applied in the twelve publications.

Monte-Carlo simulation for the theoretical  $C_{RNA}$  range. Monte-Carlo simulation for the contribution of sputum shedding on the total shedding load.

#### AUTHOR INFORMATION

##### **Corresponding Author**

\*Guangming Jiang

Email: gjiang@uow.edu.au; ORCID: 0000-0001-5399-8239

### Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

**ACKNOWLEDGMENT.** This research was supported by the Australian Research Council Discovery project (DP190100385) and a COVID-19 Digital Grant funded by Australian Academy of Science and the Department of Industry, Science, Energy and Resources through the Regional Collaborations Programme. Shuxin Zhang receives the support from a University of Wollongong PhD scholarship.

### REFERENCES

- Anand, U.; Li, X.; Sunita, K.; Lokhandwala, S.; Gautam, P.; Suresh, S.; Sarma, H.; Vellingiri, B.; Dey, A.; Bontempi, E.; Jiang, G., SARS-CoV-2 and other pathogens in municipal wastewater, landfill leachate, and solid waste: A review about virus surveillance, infectivity, and inactivation. *Environ Res* **2022**, *203*, 111839.
- Ahmed, W., Angel, N., Edson, J., Bibby, K., Bivins, A., O'Brien, J.W., Choi, P.M., Kitajima, M., Simpson, S.L., Li, J., Tschärke, B., Verhagen, R., Smith, W.J.M., Zaugg, J., Dierens, L., Hugenholtz, P., Thomas, K.V. and Mueller, J.F. 2020a. First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: A proof of concept for the wastewater surveillance of COVID-19 in the community. *The Science of the total environment* 728, 138764.
- Ahmed, W., Bertsch, P.M., Angel, N., Bibby, K., Bivins, A., Dierens, L., Edson, J., Ehret, J., Gyawali, P., Hamilton, K.A., Hosegood, I., Hugenholtz, P., Jiang, G., Kitajima, M.,

- Sichani, H.T., Shi, J., Shimko, K.M., Simpson, S.L., Smith, W.J.M., Symonds, E.M., Thomas, K.V., Verhagen, R., Zaugg, J. and Mueller, J.F. 2020b. Detection of SARS-CoV-2 RNA in commercial passenger aircraft and cruise ship wastewater: a surveillance tool for assessing the presence of COVID-19 infected travellers. *J Travel Med* 27(5).
- Ahmed, W., Bertsch, P.M., Bivins, A., Bibby, K., Farkas, K., Gathercole, A., Haramoto, E., Gyawali, P., Korajkic, A., McMinn, B.R., Mueller, J.F., Simpson, S.L., Smith, W.J.M., Symonds, E.M., Thomas, K.V., Verhagen, R. and Kitajima, M. 2020c. Comparison of virus concentration methods for the RT-qPCR-based recovery of murine hepatitis virus, a surrogate for SARS-CoV-2 from untreated wastewater. *Science of the Total Environment* 739.
- Ahmed, W.B., Paul M.; Kyle Bibby, Eiji Haramoto, Joanne Hewitt, Flavia Huygens, Pradip Gyawali, Asja Korajkic, Shane Riddell, Samendra P. Sherchan<sup>8</sup>, Stuart L. Simpson, Kwanrawee Sirikanchana, Erin M. Symonds, Rory Verhagen, Seshadri S Vasan, Masaaki Kitajima, Aaron Bivins 2020. Decay of SARS-CoV-2 and surrogate murine hepatitis virus RNA in untreated wastewater to inform application in wastewater-based epidemiology. *Environmental Research*.
- Balbi, B., Pignatti, P., Corradi, M., Baiardi, P., Bianchi, L., Brunetti, G., Radaeli, A., Moscato, G., Mutti, A. and Spanevello, A. 2007. Bronchoalveolar lavage, sputum and exhaled clinically relevant inflammatory markers: values in healthy adults. *European Respiratory Journal* 30(4), 769-781.
- Betancourt, W.Q., Schmitz, B.W., Innes, G.K., Prasek, S.M., Pogreba Brown, K.M., Stark, E.R., Foster, A.R., Sprissler, R.S., Harris, D.T., Sherchan, S.P., Gerba, C.P. and Pepper, I.L.

2021. COVID-19 containment on a college campus via wastewater-based epidemiology, targeted clinical testing and an intervention. *Science of the Total Environment* 779.
- Bivins, A., Greaves, J., Fischer, R., Yinda, K.C., Ahmed, W., Kitajima, M., Munster, V.J. and Bibby, K. 2020. Persistence of SARS-CoV-2 in Water and Wastewater. *Environmental Science & Technology Letters* 7(12), 937-942.
- Cavany, S., Bivins, A., Wu, Z., North, D., Bibby, K. and Perkins, T.A. 2021. Inferring SARS-CoV-2 RNA shedding into wastewater relative to time of infection. *medRxiv : the preprint server for health sciences*.
- CDC 2020 CDC 2019-novel coronavirus (2019-nCoV) real-time RT-PCR diagnostic pane. CDC, U. (ed).
- Chau, C.H., Strobe, J.D. and Figg, W.D. 2020. COVID-19 Clinical Diagnostics and Testing Technology. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy* 40(8), 857-868.
- Crank, K., Chen, W., Bivins, A., Lowry, S. and Bibby, K. 2022. Contribution of SARS-CoV-2 RNA shedding routes to RNA loads in wastewater. *Science of The Total Environment* 806, 150376.
- Curtis, K., Keeling, D., Yetka, K., Larson, A. and Gonzalez, R. 2021. Wastewater SARS-CoV-2 RNA Concentration and Loading Variability from Grab and 24-Hour Composite Samples. *medRxiv : the preprint server for health sciences*, 2020.2007.2010.20150607.
- EPA, E.P.A.V. 2020 Coronavirus (COVID-19): Disposing clinical waste Victoria, E.P.A. (ed), Victoria, Australia.

- Farkash, E.A., Wilson, A.M. and Jentzen, J.M. 2020. Ultrastructural Evidence for Direct Renal Infection with SARS-CoV-2. *Journal of the American Society of Nephrology* 31(8), 1683-1687.
- Fischer, C., Mögling, R., Melidou, A., Kühne, A., Oliveira-Filho, E.F., Wolff, T., Reiche, J., Broberg, E., Drosten, C. and Meijer, A. 2021. Variable sensitivity of SARS-CoV-2 molecular detection in European expert laboratories: external quality assessment, June and July 2020. *Journal of clinical microbiology* 59(3), e02676-02620.
- Gerrity, D., Papp, K., Stoker, M., Sims, A. and Frehner, W. 2021. Early-pandemic wastewater surveillance of SARS-CoV-2 in Southern Nevada: Methodology, occurrence, and incidence/prevalence considerations. *Water research X* 10, 100086.
- Gilks, W.R. 2005. Markov Chain Monte Carlo. *Encyclopedia of biostatistics* 4.
- Graham, K.E., Loeb, S.K., Wolfe, M.K., Catoe, D., Sinnott-Armstrong, N., Kim, S., Yamahara, K.M., Sassoubre, L.M., Mendoza Grijalva, L.M., Roldan-Hernandez, L., Langenfeld, K., Wigginton, K.R. and Boehm, A.B. 2021. SARS-CoV-2 RNA in Wastewater Settled Solids Is Associated with COVID-19 Cases in a Large Urban Sewershed. *Environmental Science & Technology* 55(1), 488-498.
- Hart, O.E. and Halden, R.U. 2020a. Computational analysis of SARS-CoV-2/COVID-19 surveillance by wastewater-based epidemiology locally and globally: Feasibility, economy, opportunities and challenges. *Science of The Total Environment* 730, 138875.
- Hart, O.E. and Halden, R.U. 2020b. Modeling wastewater temperature and attenuation of sewage-borne biomarkers globally. *Water Research* 172, 115473.

- Hasan, S.W., Ibrahim, Y., Daou, M., Kannout, H., Jan, N., Lopes, A., Alsafar, H. and Yousef, A.F. 2020. Detection and quantification of SARS-CoV-2 RNA in wastewater and treated effluents: Surveillance of COVID-19 epidemic in the United Arab Emirates. *Science of The Total Environment*, 142929.
- He, J., Guo, Y., Mao, R. and Zhang, J. 2021. Proportion of asymptomatic coronavirus disease 2019: A systematic review and meta-analysis. *Journal of medical virology* 93(2), 820-830.
- Huang, N., Pérez, P., Kato, T., Mikami, Y., Okuda, K., Gilmore, R.C., Conde, C.D., Gasmi, B., Stein, S. and Beach, M. 2021. SARS-CoV-2 infection of the oral cavity and saliva. *Nature medicine* 27(5), 892-903.
- Jafferli, M.H., Khatami, K., Atasoy, M., Birgersson, M., Williams, C. and Cetecioglu, Z. 2020. Benchmarking virus concentration methods for quantification of SARS-CoV-2 in raw wastewater. *The Science of the total environment*, 142939.
- Jones, D.L., Baluja, M.Q., Graham, D.W., Corbishley, A., McDonald, J.E., Malham, S.K., Hillary, L.S., Connor, T.R., Gaze, W.H., Moura, I.B., Wilcox, M.H. and Farkas, K. 2020. Shedding of SARS-CoV-2 in feces and urine and its potential role in person-to-person transmission and the environment-based spread of COVID-19. *Science of the Total Environment* 749.
- Jung, Y., Park, G.-S., Moon, J.H., Ku, K., Beak, S.-H., Lee, C.-S., Kim, S., Park, E.C., Park, D., Lee, J.-H., Byeon, C.W., Lee, J.J., Maeng, J.-S., Kim, S.-J., Kim, S.I., Kim, B.-T., Lee, M.J. and Kim, H.G. 2020. Comparative Analysis of Primer–Probe Sets for RT-qPCR of COVID-19 Causative Virus (SARS-CoV-2). *ACS Infectious Diseases* 6(9), 2513-2523.

- Kim, J.-M., Kim, H.M., Lee, E.J., Jo, H.J., Yoon, Y., Lee, N.-J., Son, J., Lee, Y.-J., Kim, M.S., Lee, Y.-P., Chae, S.-J., Park, K.R., Cho, S.-R., Park, S., Kim, S.J., Wang, E., Woo, S., Lim, A., Park, S.-J., Jang, J., Chung, Y.-S., Chin, B.S., Lee, J.-S., Lim, D., Han, M.-G. and Yoo, C.K. 2020. Detection and Isolation of SARS-CoV-2 in Serum, Urine, and Stool Specimens of COVID-19 Patients from the Republic of Korea. *Osong Public Health Res Perspect* 11(3), 112-117.
- Lai, C.-C., Liu, Y.H., Wang, C.-Y., Wang, Y.-H., Hsueh, S.-C., Yen, M.-Y., Ko, W.-C. and Hsueh, P.-R. 2020. Asymptomatic carrier state, acute respiratory disease, and pneumonia due to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): Facts and myths. *Journal of Microbiology, Immunology and Infection* 53(3), 404-412.
- Li, X., Kulandaivelu, J., Zhang, S., Shi, J., Sivakumar, M., Mueller, J., Luby, S., Ahmed, W., Coin, L. and Jiang, G. 2021a. Data-driven estimation of COVID-19 community prevalence through wastewater-based epidemiology. *Science of The Total Environment* 789, 147947.
- Li, X., Zhang, S., Shi, J., Luby, S.P. and Jiang, G. 2021b. Uncertainties in estimating SARS-CoV-2 prevalence by wastewater-based epidemiology. *Chemical Engineering Journal*, 129039.
- Lin, H.-C., Cheng, H.-F., Wang, C.-H., Liu, C.-Y., Yu, C.-T. and Kuo, H.-P. 1997. Inhaled gentamicin reduces airway neutrophil activity and mucus secretion in bronchiectasis. *American journal of respiratory and critical care medicine* 155(6), 2024-2029.

- Liu, W.-D., Chang, S.-Y., Wang, J.-T., Tsai, M.-J., Hung, C.-C., Hsu, C.-L. and Chang, S.-C. 2020. Prolonged virus shedding even after seroconversion in a patient with COVID-19. *J Infect* 81(2), 318-356.
- Matson, M.J., Yinda, C.K., Seifert, S.N., Bushmaker, T., Fischer, R.J., van Doremalen, N., Lloyd-Smith, J.O. and Munster, V.J. 2020. Effect of environmental conditions on SARS-CoV-2 stability in human nasal mucus and sputum. *Emerging infectious diseases* 26(9), 2276.
- McCall, A.-K., Palmitessa, R., Blumensaat, F., Morgenroth, E. and Ort, C. 2017. Modeling in-sewer transformations at catchment scale—implications on drug consumption estimates in wastewater-based epidemiology. *Water research* 122, 655-668.
- Miura, F., Kitajima, M. and Omori, R. 2021. Duration of SARS-CoV-2 viral shedding in faeces as a parameter for wastewater-based epidemiology: Re-analysis of patient data using a shedding dynamics model. *Science of the Total Environment* 769, 144549.
- Mohammadi, A., Esmailzadeh, E., Li, Y., Bosch, R.J. and Li, J.Z. 2020. SARS-CoV-2 detection in different respiratory sites: A systematic review and meta-analysis. *EBioMedicine* 59, 102903.
- Moreira, L.V.L., de Souza Luna, L.K., Barbosa, G.R., Perosa, A.H., Chaves, A.P.C., Conte, D.D., Carvalho, J.M.A. and Bellei, N. 2020. Test on stool samples improves the diagnosis of hospitalized patients: Detection of SARS-CoV-2 genomic and subgenomic RNA. *The Journal of infection*.
- Novazzi, F., Cassaniti, I., Piralla, A., Di Sabatino, A., Bruno, R., Baldanti, F. and San Matteo Pavia, C.-T. 2020. Detection of the SARS-CoV-2 in different biologic specimens from

- positive patients with COVID-19, in Northern Italy. *Pediatric Allergy and Immunology* 31, 72-74.
- Pan, Y., Zhang, D., Yang, P., Poon, L.L.M. and Wang, Q. 2020. Viral load of SARS-CoV-2 in clinical samples. *The Lancet Infectious Diseases* 20(4), 411-412.
- Park, C., Lee, J., Ul Hassan, Z., Ku, K.B., Kim, S.-J., Kim, H.G., Park, E.C., Park, G.-S., Park, D. and Baek, S.-H. 2021. Comparison of digital PCR and quantitative PCR with various SARS-CoV-2 primer-probe sets.
- Park, S.-k., Lee, C.-W., Park, D.-I., Woo, H.-Y., Cheong, H.S., Shin, H.C., Ahn, K., Kwon, M.-J. and Joo, E.-J. 2020. Detection of SARS-CoV-2 in fecal samples from patients with asymptomatic and mild COVID-19 in Korea. *Clinical Gastroenterology and Hepatology*.
- Pecson, B.M., Darby, E., Haas, C., Amha, Y., Bartolo, M., Danielson, R., Dearborn, Y., Di Giovanni, G., Ferguson, C. and Fevig, S. 2021. Reproducibility and sensitivity of 36 methods to quantify the SARS-CoV-2 genetic signal in raw wastewater: findings from an interlaboratory methods evaluation in the US. *Environmental Science: Water Research & Technology*.
- Peng, L., Liu, J., Xu, W., Luo, Q., Chen, D., Lei, Z., Huang, Z., Li, X., Deng, K. and Lin, B. 2020. SARS-CoV-2 can be detected in urine, blood, anal swabs, and oropharyngeal swabs specimens. *Journal of Medical Virology*.
- Pezzi, L., Charrel, R.N., Ninove, L., Nougairede, A., Molle, G., Coutard, B., Durand, G., Leparcoffart, I., de Lamballerie, X. and Thirion, L. 2020. Development and Evaluation of a duo SARS-CoV-2 RT-qPCR Assay Combining Two Assays Approved by the World

- Health Organization Targeting the Envelope and the RNA-Dependant RNA Polymerase (RdRp) Coding Regions. *Viruses* 12(6), 686.
- PICA, P.S.G. 2020. COVID-19: waste disposal for apartment living.
- Poon, K.S. and Tee, N.W.S. 2021. Detecting SARS-CoV-2 RNA in fecal specimens: The practical challenges. *Journal of Medical Virology*.
- Reese, H., Iuliano, A.D., Patel, N.N., Garg, S., Kim, L., Silk, B.J., Hall, A.J., Fry, A. and Reed, C. 2020. Estimated Incidence of Coronavirus Disease 2019 (COVID-19) Illness and Hospitalization—United States, February–September 2020. *Clinical Infectious Diseases*.
- Rose, C., Parker, A., Jefferson, B. and Cartmell, E. 2015. The characterization of feces and urine: a review of the literature to inform advanced treatment technology. *Critical reviews in environmental science and technology* 45(17), 1827-1879.
- Rusiñol, M., Martínez-Puchol, S., Forés, E., Itarte, M., Girones, R. and Bofill-Mas, S. 2020. Concentration methods for the quantification of coronavirus and other potentially pandemic enveloped virus from wastewater. *Current Opinion in Environmental Science and Health* 17, 21-28.
- Rusiñol, M., Zammit, I., Itarte, M., Forés, E., Martínez-Puchol, S., Girones, R., Borrego, C., Corominas, L. and Bofill-Mas, S. 2021. Monitoring waves of the COVID-19 pandemic: Inferences from WWTPs of different sizes. *Science of the Total Environment* 787.
- Saglietto, A., Moirano, G., Anselmino, M. and De Ferrari, G.M. 2020. Higher Testing Coverage Associated With a Lower COVID-19 Mortality Rate: Insights From Italian Regions. *Disaster Medicine and Public Health Preparedness*, 1-3.

- Saththasivam, J., El-Malah, S.S., Gomez, T.A., Jabbar, K.A., Remanan, R., Arun, K., Ogunbiyi, O., Rasool, K., Ashhab, S. and Rashkeev, S. 2021. COVID-19 (SARS-CoV-2) Outbreak Monitoring using Wastewater-Based Epidemiology in Qatar. *Science of The Total Environment*, 145608.
- Silverman, A.I. and Boehm, A.B. 2020. Systematic review and meta-analysis of the persistence and disinfection of human coronaviruses and their viral surrogates in water and wastewater. *Environmental Science & Technology Letters* 7(8), 544-553.
- Tang, A., Tong, Z.-d., Wang, H.-l., Dai, Y.-x., Li, K.-f., Liu, J.-n., Wu, W.-j., Yuan, C., Yu, M.-l., Li, P. and Yan, J.-b. 2020. Detection of Novel Coronavirus by RT-PCR in Stool Specimen from Asymptomatic Child, China. *Emerging Infectious Disease journal* 26(6), 1337.
- van Doorn, A.S., Meijer, B., Frampton, C.M., Barclay, M.L. and de Boer, N.K. 2020. Systematic review with meta-analysis: SARS-CoV-2 stool testing and the potential for faecal-oral transmission. *Alimentary pharmacology & therapeutics*.
- Walsh, K.A., Jordan, K., Clyne, B., Rohde, D., Drummond, L., Byrne, P., Ahern, S., Carty, P.G., O'Brien, K.K., O'Murchu, E., O'Neill, M., Smith, S.M., Ryan, M. and Harrington, P. 2020. SARS-CoV-2 detection, viral load and infectivity over the course of an infection. *Journal of Infection* 81(3), 357-371.
- Wang, W., Xu, Y., Gao, R., Lu, R., Han, K., Wu, G. and Tan, W. 2020. Detection of SARS-CoV-2 in different types of clinical specimens. *Jama* 323(18), 1843-1844.
- Weidhaas, J., Aanderud, Z.T., Roper, D.K., VanDerslice, J., Gaddis, E.B., Ostermiller, J., Hoffman, K., Jamal, R., Heck, P., Zhang, Y., Torgersen, K., Laan, J.V. and LaCross, N.

2021. Correlation of SARS-CoV-2 RNA in wastewater with COVID-19 disease burden in sewersheds. *Science of the Total Environment* 775.
- Wilder, M.L., Middleton, F., Larsen, D.A., Du, Q., Fenty, A., Zeng, T., Insaf, T., Kilaru, P., Collins, M., Kmush, B. and Green, H.C. 2021. Co-quantification of crAssphage increases confidence in wastewater-based epidemiology for SARS-CoV-2 in low prevalence areas. *Water Research X* 11, 100100.
- Wong, J.C.C., Tan, J., Lim, Y.X., Arivalan, S., Hapuarachchi, H.C., Mailepessov, D., Griffiths, J., Jayarajah, P., Setoh, Y.X., Tien, W.P., Low, S.L., Koo, C., Yenamandra, S.P., Kong, M., Lee, V.J.M. and Ng, L.C. 2021. Non-intrusive wastewater surveillance for monitoring of a residential building for COVID-19 cases. *The Science of the total environment* 786, 147419.
- Wu, F., Xiao, A., Zhang, J., Moniz, K., Endo, N., Armas, F., Bonneau, R., Brown, M.A., Bushman, M., Chai, P.R., Duvallet, C., Erickson, T.B., Foppe, K., Ghaeli, N., Gu, X., Hanage, W.P., Huang, K.H., Lee, W.L., Matus, M., McElroy, K.A., Nagler, J., Rhode, S.F., Santillana, M., Tucker, J.A., Wuertz, S., Zhao, S., Thompson, J. and Alm, E.J. 2020a. SARS-CoV-2 titers in wastewater foreshadow dynamics and clinical presentation of new COVID-19 cases. *medRxiv*.
- Wu, F., Zhang, J., Xiao, A., Gu, X., Lee, W.L., Armas, F., Kauffman, K., Hanage, W., Matus, M., Ghaeli, N., Endo, N., Duvallet, C., Poyet, M., Moniz, K., Washburne, A.D., Erickson, T.B., Chai, P.R., Thompson, J. and Alm, E.J. 2020b. SARS-CoV-2 Titers in Wastewater Are Higher than Expected from Clinically Confirmed Cases. *mSystems* 5(4), e00614-00620.

- Wu, Y., Guo, C., Tang, L., Hong, Z., Zhou, J., Dong, X., Yin, H., Xiao, Q., Tang, Y. and Qu, X. 2020c. Prolonged presence of SARS-CoV-2 viral RNA in faecal samples. *The lancet Gastroenterology & hepatology* 5(5), 434-435.
- Wyllie, A.L., Fournier, J., Casanovas-Massana, A., Campbell, M., Tokuyama, M., Vijayakumar, P., Warren, J.L., Geng, B., Muenker, M.C. and Moore, A.J. 2020. Saliva or nasopharyngeal swab specimens for detection of SARS-CoV-2. *New England Journal of Medicine* 383(13), 1283-1286.
- Wölfel, R., Corman, V.M., Guggemos, W., Seilmaier, M., Zange, S., Müller, M.A., Niemeyer, D., Jones, T.C., Vollmar, P., Rothe, C., Hoelscher, M., Bleicker, T., Brünink, S., Schneider, J., Ehmann, R., Zwirgmaier, K., Drosten, C. and Wendtner, C. 2020. Virological assessment of hospitalized patients with COVID-2019. *Nature* 581(7809), 465-469.
- Xiao, F., Tang, M., Zheng, X., Liu, Y., Li, X. and Shan, H. 2020a. Evidence for gastrointestinal infection of SARS-CoV-2. *Gastroenterology* 158(6), 1831-1833. e1833.
- Xiao, F., Tang, M., Zheng, X., Liu, Y., Li, X. and Shan, H. 2020b. Evidence for Gastrointestinal Infection of SARS-CoV-2. *Gastroenterology* 158(6), 1831-1833.e1833.
- Xu, Y., Li, X., Zhu, B., Liang, H., Fang, C., Gong, Y., Guo, Q., Sun, X., Zhao, D. and Shen, J. 2020. Characteristics of pediatric SARS-CoV-2 infection and potential evidence for persistent fecal viral shedding. *Nature medicine* 26(4), 502-505.
- Yang, Z., Yu, M., Li, G., Dai, X., Liu, G., Xie, J., Li, G. and Jie, Y. 2020. A Convalescent of COVID-19 with RT-PCR Test Continues Positive in Stool. *Clinical laboratory* 66(12).

- Yoon, J.G., Yoon, J., Song, J.Y., Yoon, S.Y., Lim, C.S., Seong, H., Noh, J.Y., Cheong, H.J. and Kim, W.J. 2020. Clinical significance of a high SARS-CoV-2 viral load in the Saliva. *Journal of Korean Medical Science* 35(20).
- Zhang, J., Wang, S. and Xue, Y. 2020a. Fecal specimen diagnosis 2019 novel coronavirus–infected pneumonia. *Journal of Medical Virology* 92(6), 680-682.
- Zhang, S., Li, X., Shi, J., Sivakumar, M., Luby, S., O'Brien, J. and Jiang, G. 2022. Analytical performance comparison of four SARS-CoV-2 RT-qPCR primer-probe sets for wastewater samples. *Science of The Total Environment* 806, 150572.
- Zhang, W., Du, R.-H., Li, B., Zheng, X.-S., Yang, X.-L., Hu, B., Wang, Y.-Y., Xiao, G.-F., Yan, B., Shi, Z.-L. and Zhou, P. 2020b. Molecular and serological investigation of 2019-nCoV infected patients: implication of multiple shedding routes. *Emerging Microbes & Infections* 9(1), 386-389.
- Zhao, L., Ji, Y., Yao, J., Long, S., Li, D. and Yang, Y. 2017. Quantifying the fate and risk assessment of different antibiotics during wastewater treatment using a Monte Carlo simulation. *Journal of Cleaner Production* 168, 626-631.
- Zhu, J., Guo, J., Xu, Y. and Chen, X. 2020. Viral dynamics of SARS-CoV-2 in saliva from infected patients. *Journal of Infection* 81(3), e48-e50.

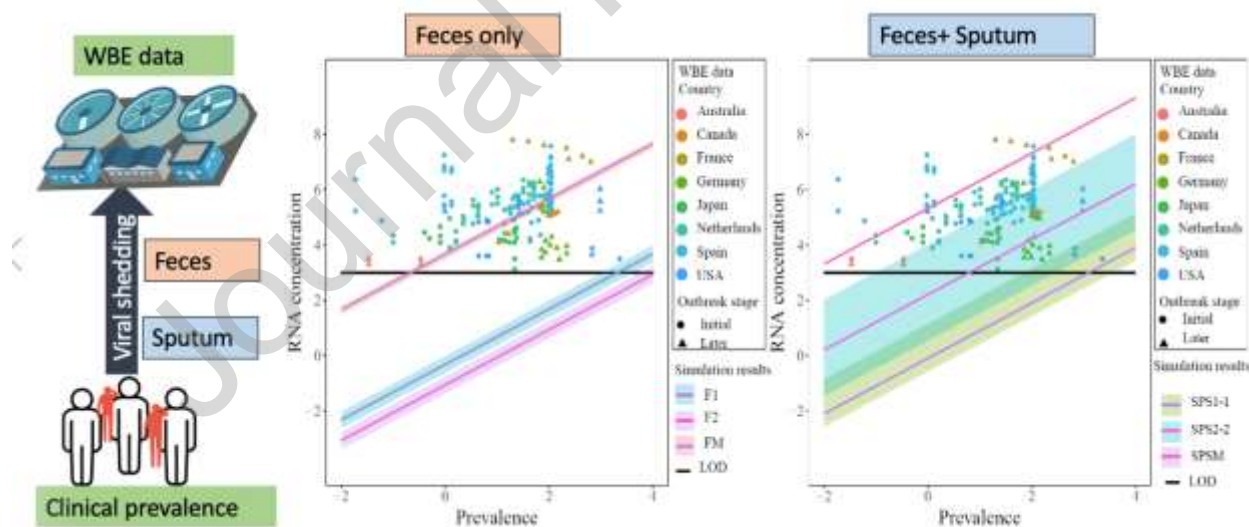
## CRediT authorship contribution statement

Xuan Li: Conceptualization, Methodology, Data curation, Writing – original draft. Jagadeeshkumar Kulandaivelu: Data curation, Writing – review & editing. Ying Guo: Data curation. Shuxin Zhang: Data curation. Jiahua Shi: Data curation. Jake O’Brien: Writing – review & editing. Sudipti Arora: Writing – review & editing. Manish Kumar: Writing – review & editing. Samendra P. Sherchan: Writing – review & editing. Ryo Honda: Writing – review & editing. Greg Jackson: Writing – review & editing. Stephen P. Luby: Writing – review & editing. Guangming Jiang: Conceptualization, Methodology, Writing – review & editing, Funding acquisition.

## Declaration of Competing Interest

☑ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Graphical abstract



## Environmental Implication

The contribution of SARS-CoV-2 RNA shedding sources in sewers is critical for the accuracy of COVID-19 surveillance through wastewater-based epidemiology (WBE) approach, where fecal shedding was considered as the major source. This study suggests that sputum is potentially

a critical shedding source to wastewater other than feces. Sputum discharge behavior also led to great variations for the SARS-CoV-2 RNA in wastewater, which potentially affect the accuracy of WBE surveillance for COVID-19 infections. Thus, sputum and its discharging behavior from patients are recommended to be considered in future wastewater surveillance and studies about the viral circulation in the environment.

### **Highlights**

- A review of SARS-CoV-2 shedding sources and their parameters in wastewater.
- Major SARS-CoV-2 shedding sources (feces and sputum) were assessed by Monte Carlo simulations.
- Feces, as a major shedding source, contribute partly to SARS-CoV-2 in wastewater.
- Sputum is confirmed by simulations as a major source of SARS-CoV-2 in wastewater.